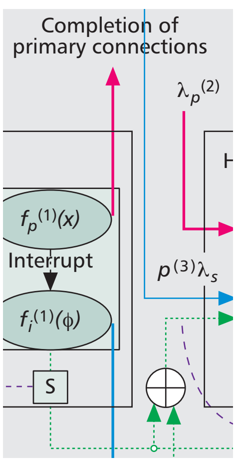
ACCEPTED FROM OPEN CALL

A QUEUEING-THEORETICAL FRAMEWORK FOR QOS-ENHANCED SPECTRUM MANAGEMENT IN COGNITIVE RADIO NETWORKS

LI-CHUN WANG, CHUNG-WEI WANG, AND KAI-TEN FENG, NATIONAL CHIAO TUNG UNIVERSITY



The authors outline the fundamental modeling issues of opportunistic spectrum access in cognitive radio networks. They identify the effects of connection-based channel usage on the QoS performance of spectrum management techniques.

ABSTRACT

This article outlines the fundamental modeling issues of opportunistic spectrum access in cognitive radio networks. In particular, we identify the effects of connection-based channel usage on the QoS performance of spectrum management techniques. During the transmission period of a secondary user's connection, the phenomenon of multiple spectrum handoffs due to interruptions of primary users arises quite often. In addition to multiple interruptions, the connection-based channel usage behaviors are also affected by spectrum sensing time, switching between different channels, generally distributed service time, and channel contention between multiple secondary users. An analytical framework based on the preemptive resumption priority M/G/1 queueing theory is introduced to characterize the effects of the above factors simultaneously. The proposed generalized analytical framework can incorporate various system parameters into the design of very broad spectrum management techniques, including spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility. The applications of this analytical framework on spectrum decision as well as spectrum mobility are discussed, and some open issues using this framework are suggested in this article.

INTRODUCTION

Cognitive radio (CR) allows low-priority secondary users to temporarily utilize the unused licensed channel of high-priority primary users, thereby significantly improving overall spectrum efficiency [1]. However, the secondary users need to vacate the occupied channel when the primary users appear. In order to return the occupied channel to the primary user and discover a suitable target channel to resume the unfinished transmission, *spectrum handoff* procedures are initiated for the secondary users.[1] Basically, according to the principle of selecting the target channel for spectrum handoff, the operating modes of the secondary networks can be categorized as non-hopping and hopping. In the non-hopping mode, the secondary user always stays on its current operating channel when it is interrupted, which is the basic mode of IEEE 802.22 systems [3]. In the hopping mode, the interrupted secondary user can stay on its current operating channel or change to another channel according to traffic statistics. An example of the hopping mode is the phase-shifting hopping method used in IEEE 802.22 systems [3]. Secondary users' connections may execute multiple handoffs during its transmission period due to interruptions from primary users [4]. Clearly, these handoffs will degrade the quality of service (QoS) performance of the secondary users in providing sensitive traffic.

In order to overcome the performance degradation due to multiple spectrum handoffs in the non-hopping or hopping mode, various spectrum management techniques are re-examined from the perspective of link connection quality for the secondary users. The spectrum management techniques include:

- Spectrum sensing (detecting an unused channel in the sensing phase)

This work was presented in part at the IEEE International Performance Computing and Communications Conference (IPCCC), 2008. This work was supported in part by the MoE ATU Plan and the National Science Council (NSC) grant 97-2917-I-009-109.

[1] *Spectrum handoff in CR networks is different from the conventional handoff mechanisms in cellular mobile networks. Spectrum handoff considers two types of users with different priorities, where the high-priority primary users have the right to interrupt the transmission of the low-priority secondary users. When the interruption event occurs, the secondary user must stop using the current channel even though the received signal strength is still acceptable. In contrast, all users in the conventional handoff mechanisms have the same priority to access channels, and they change their operating channels mainly due to deterioration of signal quality [2].*

1536-1284/11/$25.00 © 2011 IEEE

- Spectrum decision (selecting the best initial channel in the analysis phase)
- Spectrum sharing (avoiding collision in the access phase)
- Spectrum mobility (switching to a suitable channel when a licensed primary user appears)

Referring to [5], the relationships of these four spectrum management functionalities are shown in Fig. 1. In the figure, the secondary users first request channels from the CR network. With the spectrum decision functionality, they can determine their initial operating channels from all M candidate channels based on the spectrum sensing outcomes. The spectrum sharing functionality must be implemented to alleviative channel contention when multiple secondary users access the same channel. Furthermore, if a primary user appears on the occupied channel, the spectrum handoff procedures in the spectrum mobility functionality must be initiated.

In this article, in order to evaluate the QoS performance of spectrum management techniques in the hopping mode, an analytical framework based on the preemptive resumption priority (PRP) M/G/1 queueing theory is developed. The proposed analytical framework can provide important insights into the design of the system parameters of spectrum management techniques for various traffic arrival rates and service time distributions. The effectiveness of the proposed analytical framework is illustrated by some examples. Specifically, we investigate how to design a load balancing spectrum decision scheme and evaluate the latency performance of various spectrum handoff schemes based on the proposed analytical framework. In conclusion, the transmission latency of secondary users can be improved significantly if they can adaptively adopt the optimal system parameters according to traffic conditions. Finally, we suggest some open issues on top of the proposed analytical model.

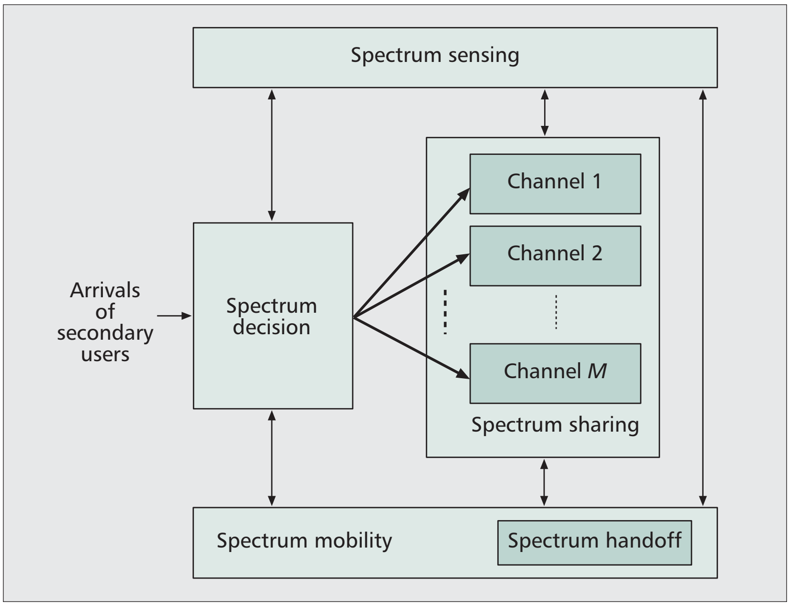


Figure 1. *Relationships between spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility functionalities.*

Design Features and Current Solutions

Design Features

An important issue for a CR network is to develop an analytical framework to characterize the behaviors of the connection-based channel usage, including:
- Multiple interruptions and handoffs
- Spectrum sensing time
- Various operating channels
- Generally distributed service time
- Channel waiting time due to multiple secondary users

Many analytical models have been proposed to characterize these features [6–9]. However, these five key design features have not been considered simultaneously.

Current Solutions

As shown in Table 1, the current modeling techniques of channel usage behaviors in CR networks can be classified into three categories: the partially observable Markov decision process (POMDP), the two-dimensional Markov chain, and the PRP M/G/1 queueing model.

In [6], the evolution of the channel usage of the primary network is characterized by a discrete-time Markov chain that has two occupancy states (busy and idle). The framework of POMDP was developed to preselect the best action (target channel) to maximize the immediate reward (expected per-slot throughput) of the decision maker (secondary user) at the next time slot [6]. In addition to the effects of the traffic loads of the primary network, the effects of the traffic loads of both the primary and secondary users on the statistics of channel occupancy were considered in [7], where a two-dimensional Markov chain is used to represent the total numbers of primary and secondary users in a CR network, respectively. When the secondary users are interrupted, it is assumed that they can immediately find an idle channel if at least one idle channel exists. As a result, the spectrum sensing time is neglected in this model. Furthermore, the Markov chain model is more suitable for exponentially distributed service time. It is unclear how to extend this kind of model to the case with generally distributed service time.

Some researchers have used the PRP M/G/1 queueing model to characterize the spectrum usage behaviors of each channel. Based on this model, the effects of multi-user contention and multiple interruptions on the latency performance of secondary users' connections were studied in [8, 9]. However, this PRP M/G/1 queueing model considered the non-hopping mode; thus, there is only one candidate channel for spectrum handoff. Note that the sensing time issue is not addressed in this model.

Challenges

Compared to the traditional PRP M/G/1 queueing model, the proposed PRP M/G/1 queueing network model resolves two new challenging conditions:
- Various operating channels in the hopping model
- Spectrum sensing time

Model Name	Multiple spectrum handoffs	Spectrum sensing time	Various operating channels	General service time	Multiple secondary connections
POMDP [6]	○	×	○	×	×
Two-dimensional Markov chain [7]	○	×	○	×	○
PRP M/G/1 queuing model [8, 9]	○	×	×	○	○
Proposed RPR M/G/1 queuing network model	○	○	○	○	○

Table 1. *Comparison of various analytical models for CR Networks, where the signs "○" and "×" indicate that the proposed model "does" and "does not" consider the corresponding feature, respectively.*

First, it is challenging to find a unifying model to characterize the channel switching behaviors for various target channel selection methods. We provide a systematic approach based on the proposed PRP M/G/1 queueing network model to catch the randomness property of the target channel selection, and evaluate its effects on the system performance metrics of transmission latency and channel utilization. Second, in the hopping mode secondary users may need to perform spectrum sensing to search for idle channels. The suggested PRP M/G/1 queueing network can also easily incorporate the sensing time into performance analysis.

Transmission Processes with Multiple Handoffs for the Secondary Users' Connections

Figure 2 illustrates the transmission processes of a secondary connection in a two-channel CR network. The procedures consist of the following steps:

- In Fig. 2a, a secondary user plans to establish a new connection flow SC_A to its intended receiver.
- Next, in Fig. 2b, the transmitter and receiver decide on their initial operating channel for SC_A. In this example, they can select channel Ch1 or Ch2.
- In Fig. 2c, SC_A is established at Ch1. During the transmission period of SC_A, a request from a primary connection may arrive at Ch1.
- Next, in Fig. 2d, SC_A detects the appearance of a primary user.[2] Then the spectrum handoff procedures are initiated to vacate Ch1 and discover a suitable target channel to resume the unfinished transmission.
- In Fig. 2e, the target channel of SC_A is decided for spectrum handoff. If the non-hopping mode is adopted, the operating channel of SC_A cannot be changed; thus, SC_A must select Ch1 to be its target channel. However, SC_A can select Ch1 or Ch2 for its target channel when the hopping mode is adopted. There are many methods of selecting the target channel. For example, the target channel can be searched for by instantaneous spectrum sensing at this moment of interruption. In this case, the effect of spectrum sensing time τ on the latency performance of SC_A must be considered.
- Finally, if SC_A chooses to stay on Ch1, its remaining transmission will be resumed after all traffic loads of the primary users at Ch1 have been served, as shown in Fig. 2f. On the other hand, if the decision is to change to Ch2, there are two possible situations. If Ch2 is idle, SC_A can transmit remaining data immediately as shown in Fig. 2g. Otherwise, SC_A must wait at the queue until all secondary users in the present queue of Ch2 are served, as shown in Fig. 2h.
- Note that similar spectrum handoff behaviors may be executed many times because a secondary user's connection may experience multiple interruptions from primary users during its transmission period. Hence, the procedures from Figs. 2c–2h will be executed repeatedly. In this article, a set of target channels, called the *target channel sequence*, will be selected sequentially.

Queueing-Theoretical Framework for Spectrum Management

Overview of PRP M/G/1 Queueing Network Model

Now we propose a preemptive resume priority (PRP) M/G/1 queueing network model to characterize the connection-based spectrum usage behaviors in CR networks. This model is very general and can easily be adjusted to evaluate the performance of various spectrum management techniques under different traffic conditions. Furthermore, it can also be applied to general CR network architectures, including ad hoc and centralized CR networks. Key features of the proposed queueing network model are listed below:

- Each server (channel) has two types of customers (connections). Before transmitting data, the traffic of the primary and secondary users enter the high-priority and low-priority queues,[3] respectively. According to the traffic arrival time at queues, the *primary* and *secondary connections* can be established without any collisions. The first come first served (FCFS) scheduling discipline is adopted to arrange the transmission order for these connections with the same priority.[4]

---

[2] *We assume that the considered CR network is a time-slotted system. In order to detect the presence of primary users, each secondary user must perform spectrum sensing at the beginning of each time slot. If the current operating channel is idle, the secondary user can transmit one slot-sized frame in this time slot. Otherwise, the secondary user must perform spectrum handoff procedures to resume its unfinished transmission on the target channel. Furthermore, the secondary user can differentiate the appearance of a primary user or secondary user by existing spectrum sensing techniques such as feature detection. The issue of differentiating primary and secondary users is beyond the scope of this article.*

[3] *Note that we assume the considered two queues to have infinite length.*

[4] *In fact, the analytical results of mean values obtained from the proposed framework can be applied to another scheduling discipline that is independent of the service time of the primary and secondary connections because the averages of system performance metrics will be invariant to the order of service in this case [10, p. 113].*

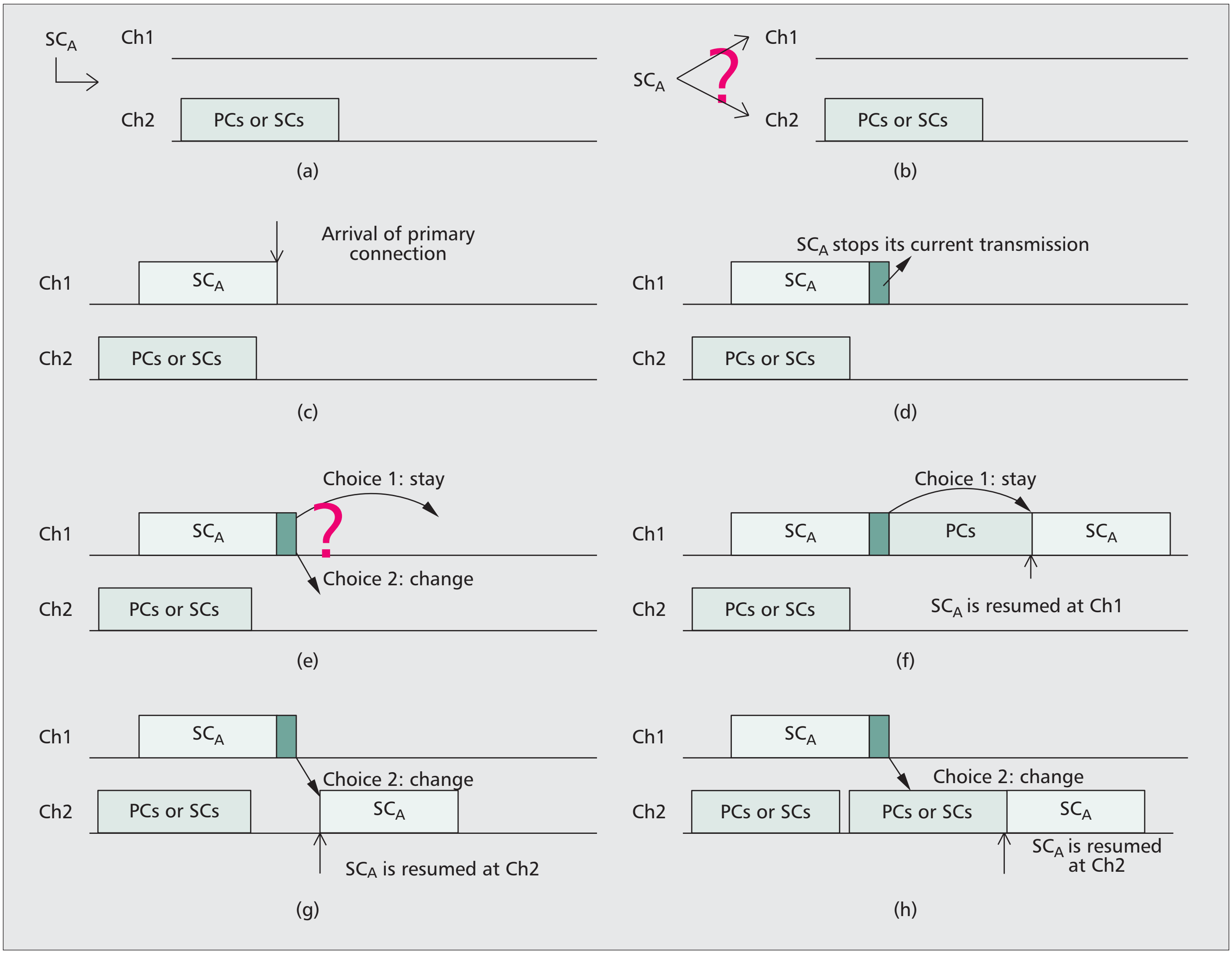


Figure 2. *Illustration of transmission procedures in a two-channel system. The medium green areas indicate that the channels are occupied by existing primary users' connections (PCs) or other secondary users' connections (SCs): a) the transmitter of the secondary connection SC_A plans to establish a connection flow to the intend receiver; b) the transmitter and receiver of SC_A can select channel Ch1 or Ch2 as the initial operating channel; c) during the transmission period of SC_A, a primary connection arrives at Ch1; d) the transmission of SC_A is stopped; e) the target channel of SC_A is decided for spectrum handoff, they can either stay on Ch1 or change to Ch2; f) SC_A vacates Ch1 and then resumes the unfinished transmission when Ch1 becomes idle; g) SC_A vacates Ch1 and changes its operating channel to the idle channel, Ch2; h) SC_A vacates Ch1 and changes its operating channel to the busy channel, Ch2.*

- Primary users have the preemptive priority to interrupt the transmissions of secondary users. Interrupted secondary users can resume the unfinished transmission on the selected target channel instead of all the data.
- A secondary connection may experience multiple interruptions from primary users during its transmission period.

Note that this model can be also extended to characterize the effects of heterogeneous channel bandwidth [11]. Some assumptions are adopted for ease of analysis.

- The arrival processes of the primary and secondary connections are Poisson.
- Only one user can transmit on each channel at any time instant.
- The secondary transmitter can notify its corresponding receiver with the interruption event according to spectrum handoff protocols [12].

Figure 3 shows an example of the PRP M/G/1 queueing network model with three channels. Let λ_s be the arrival rate of the secondary connections in a CR network. When a secondary connection arrives at the CR network, it can select its initial operating channel. Let $p^{(k)}$ be the probability of selecting channel k. Thus, the effective arrival rate of the secondary connection at channel k is $\lambda_s^{(k)} = p^{(k)} \lambda_s$. Note that various spectrum decision algorithms will result in different values of $p^{(k)}$.

When a newly arriving secondary user's data is connected to the low-priority queue of its idle initial operating channel, it can be transmitted immediately. If a primary connection appears at channel k, a secondary connection using channel k will be interrupted. In this case, the secondary connection can decide to stay on the current operating channel or change to another channel through different feedback paths, depending on the operating mode and the adopted spectrum handoff scheme. If choosing to stay on its current operating channel, the interrupted secondary

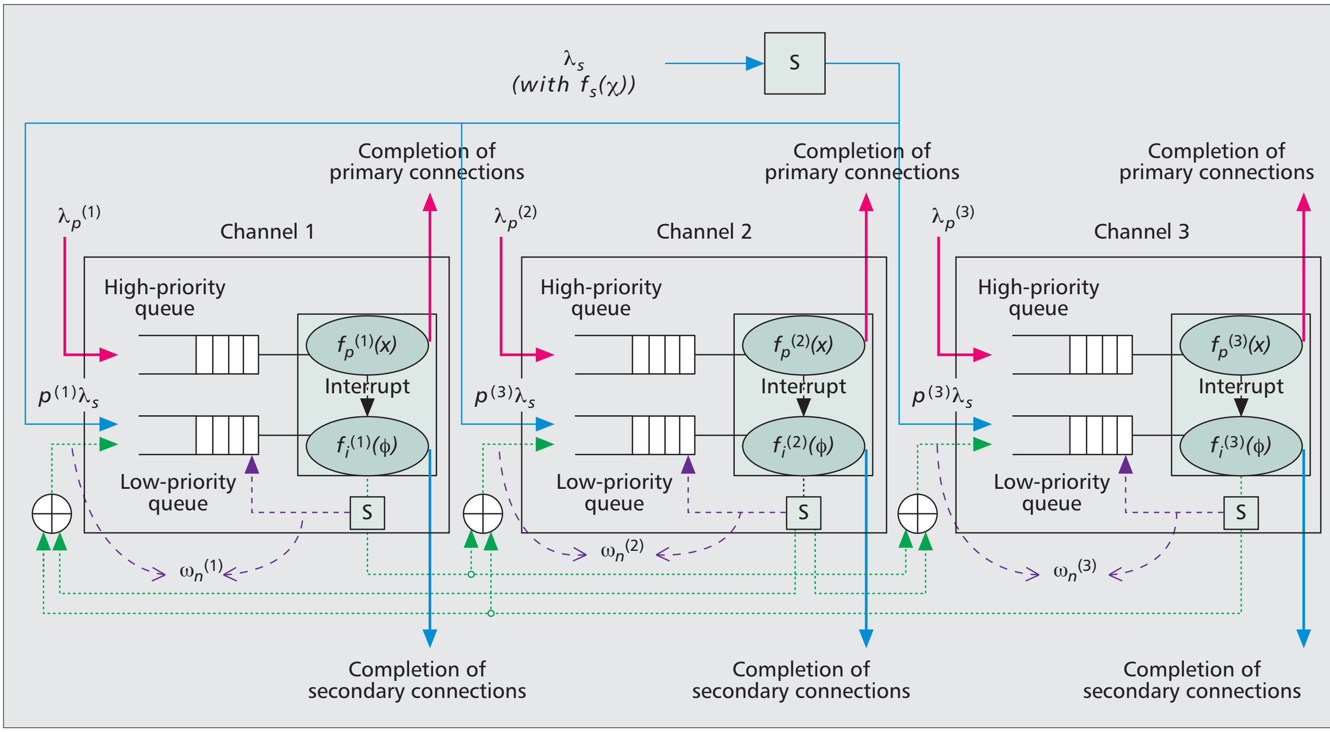


Figure 3. *The PRP M/G/1 queueing network model with three channels.* $\lambda_p^{(k)}$, $\lambda_s^{(k)}$, *and* $\omega_n^{(k)}$ *are the arrival rates of the primary connections, the secondary connections, and the type-*n *secondary connections (*$n \geq 1$*) at channel* k*. Note that* $\omega_0^{(k)} = \lambda_s^{(k)}$*. Furthermore,* $f_p^{(k)}(x)$ *and* $f_i^{(k)}(\phi)$ *are the PDFs of* $X_p^{(k)}$ *and* $\Phi_i^{(k)}$*, respectively.*

connection places the remaining data and wait at the head of the low-priority queue of the current operating channel. If it decides to change its operating channel, its remaining data will be connected to the tail of the low-priority queue of another channel. Note that ⊕ indicates that the traffic loads of the interrupted secondary connections are merged. When the interrupted secondary connection transmits its remaining data on the selected target channel, it might be interrupted again. Hence, the proposed model can describe the effects of multiple handoffs.

In Fig. 3, **S** represents the channel selection point, where the newly arriving secondary connection must select its initial operating channel, or the interrupted secondary connection must select its target channel for spectrum handoff. There are many methods to select these channels. For example, the secondary connection can decide its initial operating channel or target channel according to the predetermined probability or the outcomes from instantaneous spectrum sensing. If the spectrum sensing is executed to search the idle channels, **S** can be regarded as a tapped delay line or a server with constant service time, which is related to sensing time. Hence, the effects of spectrum sensing time on the latency performance of secondary connections can be characterized.

Modeling of Connection-based Channel Usage Behaviors

Now, we explain why the proposed model can characterize the connection-based channel usage behaviors of a CR network. In order to accurately characterize the transmission processes of a secondary connection, we must take the seven events discussed earlier into account.

- Secondary connection arrival event (Fig. 2a): We assume that the arrival process of the secondary connections is Poisson. Let X_s and $f_s(x)$ be the service time of the secondary connections and its probability density function (PDF), respectively.
- Initial channel selection event of the secondary connections (Fig. 2b): We use $p^{(k)}$ to represent the probability that the secondary connection selects channel k for its initial operating channel. Furthermore, if spectrum sensing is executed to decide the initial operating channel, the effect of sensing time can be modeled by **S**.
- Primary connection arrival event (Fig. 2c): We assume that the arrival process of the primary connection is Poisson. Denoteby $\lambda_p^{(k)}$ the arrival rate of the primary connections with default channel k. Furthermore, let $X_p^{(k)}$ and $f_p^{(k)}(x)$ be the service time of the primary connection at channel k and its corresponding PDF.
- Interruption event (Fig. 2d): In the PRP M/G/1 queueing network model, the primary users have the preemptive priority and can interrupt the secondary user's transmission. In other words, the secondary users must vacate the occupied channel when the primary users appear.
- Target channel selection event (Fig. 2e): An interrupted secondary connection can stay on its current channel or change to another channel. To this end, its remaining transmission must be connected to the low-priority queue of the current channel or another channel through different feedback paths. Furthermore, if the spectrum sensing is executed to search the target channel, the effects of sensing time can be modeled by **S**.

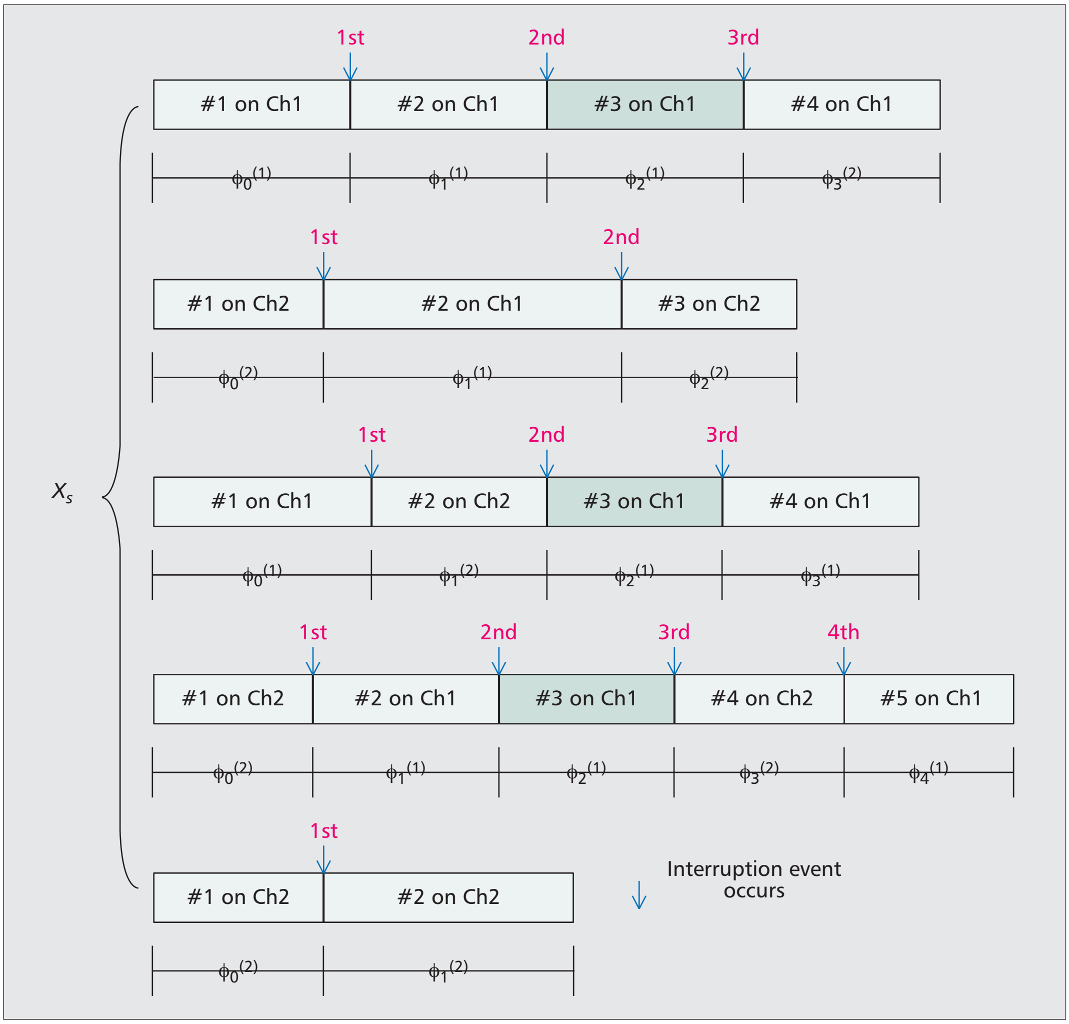


Figure 4. *Illustration of the physical meaning of random variable $\Phi_i^{(k)}$. For example, $\Phi_2^{(1)}$ is one of the third segments (darker areas) of the first, third, and fourth secondary connections.*

When the interrupted secondary connection transmits its remaining data on the selected target channel, it is possible to be interrupted again. Hence, the proposed model can describe the effects of multiple handoffs.

- Resumption event (Figs. 2f–h): The interrupted secondary connection can resume its unfinished transmission on the target channel instead of retransmitting all of the data.
- Multiple handoff events: Two auxiliary parameters ($\omega_i^{(k)}$ and $\Phi_i^{(k)}$) are suggested to characterize the traffic flows of interrupted secondary connections.

Two Auxiliary Parameters: $\omega_i^{(k)}$ and $\Phi_i^{(k)}$

In Fig. 3, we use two auxiliary parameters to characterize the traffic flows of the interrupted secondary connections. Type-*i* secondary connections represent secondary connections that have experienced *i* interruptions. Denote $\omega_i^{(k)}$ as the arrival rate of traffic flows redirected from the type-(i – 1) secondary connections at channel k. That is, $\omega_i^{(k)}$ is the arrival rate of the type-*i* secondary connections at channel k. Note that $\omega_0^{(k)} = \lambda_s^{(k)}$. Furthermore, let $\Phi_i^{(k)}$ be the transmission duration of a secondary connection between the ith and the $(i+1)^{th}$ interruptions at channel k and $f_i^{(k)}(\Phi)$ be the PDF of $\Phi_i^{(k)}$. That is, $\Phi_i^{(k)}$ is the effective service time of the type-*i* secondary connections at channel k.

Figure 4 illustrates the physical meaning of random variable $\Phi_i^{(k)}$. Recall that X_s is the service time of the secondary connections. We generate X_s five times in Fig. 4, where the five realizations are divided into many segments due to multiple primary users' interruptions. For example, the first secondary connection (realization) is divided into four segments because it experiences three interruptions at channels 1, 1, 1, and 2, respectively. Thus, this secondary connection's initial operating channel is Ch1, and its target channel sequence is (Ch1, Ch1, Ch2). In Fig. 4, random variable $\Phi_2^{(1)}$ is one of the darker regions, representing the transmission duration of a secondary connection between the second and third interruptions at Ch1. That is, $\Phi_2^{(1)}$ is one of the third segments of the first, third, and fourth secondary connections in Fig. 4. Note that the fifth secondary connection in Fig. 4 does not have a third segment because it is interrupted only once.

In the hopping mode, it is quite complex to find the PDF of the effective service time of each segment because the effective service time is dependent on the traffic statistics of the primary and other secondary users of each channel, and the operating channels for these segments can be different. Based on the proposed analytical framework, we provide a systematic approach to evaluate the impacts of various system param-

An interesting open problem is to determine which spectrum handoff scheme can result in the shortest extended data delivery time under various traffic parameters and sensing time.

eters on the effective service time and can help derive the closed-form expression for the PDF of the effective service time of each segment.

Constraints

Finally, we denote $\rho^{(k)}$ as the busy probability of channel k. In an M-channel network, the following constraint shall be satisfied:

$$\rho^{(k)} \triangleq \lambda_p^{(k)} \mathbf{E}\left[X_p^{(k)}\right] + \sum_{i=0}^{\infty} \omega_i^{(k)} \mathbf{E}\left[\Phi_i^{(k)}\right] < 1, \qquad (1)$$

where $\rho^{(k)}$ can be also interpreted as the utilization factor of channel k.

In the following, we provide two examples to illustrate the effectiveness of the proposed analytical framework. Specifically, we focus on designing the load balancing spectrum decision scheme and the evaluation of latency performance for various spectrum handoff schemes.

QoS Issues in Spectrum Decision

Spectrum decision helps the secondary user select the best initial channel to transmit data. In order to evenly distribute the traffic loads of secondary users to candidate channels, an effective spectrum decision scheme should take into account not only the traffic statistics of the primary and secondary users, but also interruptions from the primary users.

Here, we focus on the probability-based channel selection scheme, where the secondary user selects its initial operating channel from M candidate channels based on the predetermined distribution probability vector (denoted by $\boldsymbol{p} = (p^{(1)}, p^{(2)}, º, p^{(M)})$). Obviously, this scheme must consider the load balancing issues and prevent the secondary connections from selecting a busy channel. Hence, it is important to determine the optimal channel selection probability to minimize the transmission latency of the secondary connections. To this end, we formulate an optimization problem to find the optimal distribution probability vector (denoted by $\boldsymbol{p}^*$) to minimize the average overall system time of the secondary connections (denoted by $\mathbf{E}[S]$), which is defined as the duration from the instant data arrive at the system until the instant the whole transmission is finished. Specifically,

$$\boldsymbol{p}^* = \underset{\forall\ \boldsymbol{p}}{\arg\min}\, \mathbf{E}\left[S(\boldsymbol{p})\right]. \qquad (2)$$

The closed-form expression for $\mathbf{E}[S]$ was derived in [11] based on the proposed model. Hence, $\boldsymbol{p}^*$ can be determined easily.

QoS Issues in Spectrum Mobility

Target channel selection in spectrum mobility functionality is an important problem. Unlike the traditional PRP M/G/1 queueing model considering only the non-hopping mode, we further consider the hopping mode in our proposed PRP M/G/1 queueing network model. According to the decision timing for selecting target channels, handoff mechanisms in the hopping mode can be categorized into proactive and reactive handoff schemes [13, 14]. The proactive handoff scheme *proactively* determines the target channel sequence *before* data transmission. In this case, it is necessary to resolve the issue of channel obsolescence because the preselected target channel may not be available when the handoff is requested. Because the channel obsolescence issue will increase the extended data delivery time, the key challenge is to predetermine the optimal target channels to minimize the extended data delivery time, especially in the case of multiple handoffs. Here, the extended data delivery time of a secondary connection is defined as the duration from the instant of transmitting data until the instant of finishing the whole transmission. On the other hand, for the reactive handoff scheme, the target channel is searched for by *reactively* spectrum sensing *after* handoff request is initiated. Then the secondary users can resume the unfinished transmission on one of the idle channels. Hence, the target channel sequence is a random sequence depending on the traffic statistics and sensing outcomes. In this case, spectrum sensing time is a key dominant factor for the extended data delivery time.

Both proactive and reactive handoff schemes have their own advantages and disadvantages. The reactive handoff scheme can reduce its extended data delivery time by avoiding selecting a busy channel, but at the cost of relying on fast spectrum sensing techniques. The proactive handoff can save the time of scanning the whole spectrum to determine the target channel, but the predetermined target channel may not be available when it is requested by the secondary user. Thus, an interesting open problem is to determine which spectrum handoff scheme can result in the shortest extended data delivery time under various traffic parameters and sensing times [13, 14].

Figure 5 compares the extended data delivery time for the proactive and reactive handoff schemes. Here, we consider a two-channel system with the following traffic parameters: $\lambda_s^{(1)} = \lambda_s^{(2)} = 0.01$, $\mathbf{E}[X_p^{(1)}] = \mathbf{E}[X_p^{(2)}] = 10$, and $\lambda_p^{(2)} = \lambda_p^{(2)} = \lambda_p$. From this figure, we have the following important observations. First, the extended data delivery time of the reactive handoff has a singular point at $\lambda_p = 0.043$. This is because the two different predetermined target channel sequences are adopted in the cases of $\lambda_p < 0.043$ and $\lambda_p > 0.043$. Based on the proposed model, the traffic-adaptive proactive handoff scheme can be designed to appropriately change to a better target channel sequence according to traffic conditions. Next, we focus on the reactive handoff scheme. In the ideal case where spectrum sensing time (denoted by τ) is 0, the extended data delivery time can be shortened around 7~20 percent compared to the proactive handoff scheme over various arrival rates of the primary connections. This is because the reactive handoff scheme can perform spectrum sensing to find the idle channels immediately. When $\tau = 5$, the reactive handoff scheme is not always better than the proactive handoff scheme. Specifically, when $\lambda_p < 0.037$, the proactive handoff scheme can result in shorter extended data delivery time. Hence, we can conclude that the proactive handoff scheme yields shorter extended data delivery time than the reactive handoff scheme when the traffic loads of the primary users are light, whereas the reactive scheme performs better with heavy traffic loads. Finally, the reactive handoff scheme will result in the longest extended data delivery time when τ = 10. Based on the proposed model, a principle to determine which handoff scheme should be adopted in CR networks can be designed for various sensing times and traffic parameters.

Open Issues

The proposed queueing network model provides a systematic method to help the design of spectrum management technologies. It can catch the five general behaviors for connection-based channel usage. More applications have been discussed in [11, 13–15]. Some interesting research issues that can be extended from the proposed model include the following:

- An important research direction of *spectrum sensing* is to consider the effects of missed detection and false alarm on the latency performance of primary and secondary users. These two kinds of sensing errors lead to the extension of the overall system time of primary and secondary connections. Preliminary results using the proposed analytical model with imperfect sensing can be found in [11].
- From the viewpoint of *spectrum decision*, it is worthwhile to determine the optimal distribution probability vector for the probability-based spectrum decision method when secondary connections may have different opinions on the observed traffic statistics $\lambda_p^{(k)}$, λ_s, $f_p^{(k)}(x)$, and $f_s(x)$.
- An interesting *spectrum sharing* issue is to incorporate the distributed channel contention behaviors into the proposed model. In the proposed model, we assume that the FCFS scheduling policy is adopted. For a distributed medium access control (MAC) protocol such as carrier sense multiple access (CSMA), the channel contention time and retransmission in the MAC layer should be taken into account when calculating the latency performance of secondary users. Furthermore, the proposed queueing network model can facilitate the design of an interference-avoiding admission control mechanism, discussed briefly in [15].
- The proposed model assumes that the *spectrum mobility* functionality can help an interrupted secondary user resume its unfinished data transmission on a suitable channel. This resumption policy can be characterized by the preemptive resumption priority queueing network. However, in other scenarios, an interrupted secondary user may need to retransmit the whole connection rather than resuming the unfinished transmission. In this situation, a CR network should be modeled by the preemptive repeat priority queueing network. It is also worthwhile to investigate the latency performance results from different transmission policies.

Conclusions

In this article, the preemptive resume priority (PRP) M/G/1 queueing network model has been proposed to evaluate the QoS performance of various spectrum management techniques in the non-hopping and hopping modes. This analytical framework can characterize the general behavior of connection-based channel usage and help evaluate QoS performance in various traffic conditions. In order to demonstrate the effectiveness of this model, we present two examples to explain how to design system parameters for the probability-based spectrum decision schemes and evaluate the latency performance of different spectrum handoff schemes. There are still many open problems for spectrum management techniques. On top of the proposed analytical framework, it is expected that better traffic-adaptive solutions can be provided to solve these open issues from a systematic viewpoint.

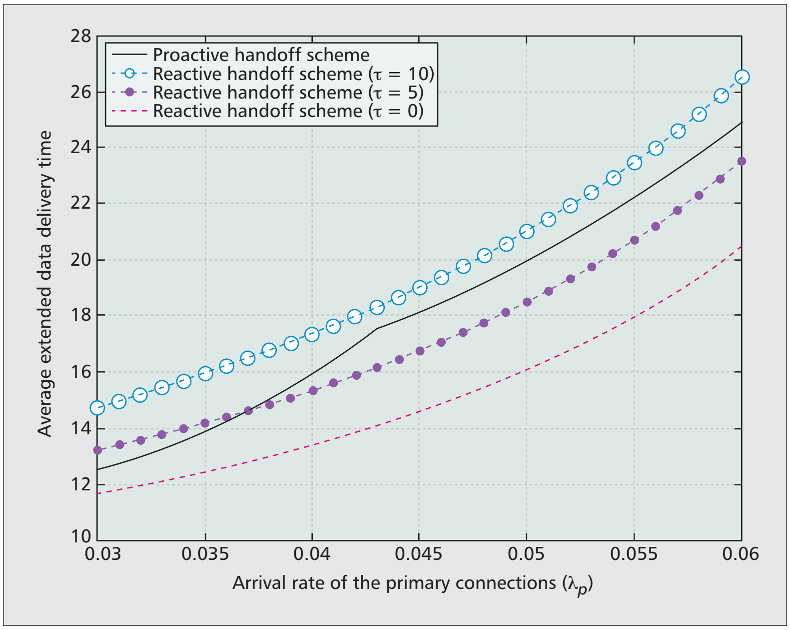


Figure 5. *Comparison of the average extended data delivery time for different spectrum handoff schemes, where* $\mathbf{E}[X_s] = 10$.

References

[1] I. F. Akyildiz *et al.*, "A Survey on Spectrum Management in Cognitive Radio Networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, Apr. 2008.
[2] A. Sgora and D. D. Vergados, "Handoff Prioritization and Decision Schemes in Wireless Cellular Networks: A Survey," *IEEE Commun. Surveys and Tutorials*, vol. 11, no. 4, 2009, pp. 57–77.
[3] W. Hu *et al.*, "Dynamic Frequency Hopping Communities for Efficient IEEE 802.22 Operation," *IEEE Commun. Mag.*, vol. 45, no. 5, May 2007, pp. 80–87.
[4] H.-J. Liu *et al.*, "Study on the Performance of Spectrum Mobility in Cognitive Wireless Network," *IEEE Singapore Int'l. Conf. Commun. Sys.*, June 2008.
[5] C.-W. Wang, *Queueing-Theoretical Spectrum Management Techniques for Cognitive Radio Networks*, Ph.D. dissertation, Nat'l. Chiao-Tung Univ., Sept. 2010.
[6] Q. Zhao and A. Swami, "A Decision-theoretic Framework for Opportunistic Spectrum Access," *IEEE Wireless Commun.*, vol. 14, no. 4, Aug. 2007, pp. 14–20.
[7] X. Zhu, L. Shen, and T.-S. P. Yum, "Analysis of Cognitive Radio Spectrum Access with Optimal Channel Reservation," *IEEE Commun. Letters*, vol. 11, no. 4, Apr. 2007, pp. 304–06.
[8] C. Zhang, X. Wang, and J. Li, "Cooperative Cognitive Radio with Priority Queueing Analysis," *IEEE ICC '09*, June 2009.
[9] P. Zhu, J. Li, and X. Wang, "Scheduling Model for Cognitive Radio," *Int'l. Conf. Cognitive Radio Oriented Wireless Networks and Commun.*, May 2008.
[10] L. Kleinrock, *Queueing Systems — Volume 2: Computer Applications*, Wiley, 1975.

[11] L.-C. Wang, C.-W. Wang, and F. Adachi, "Load-Balancing Spectrum Decision for Cognitive Radio Networks," *IEEE JSAC*, vol. 29, no. 4, Apr. 2011, pp. 757–69.
[12] X. Liu and Z. Ding, "ESCAPE: A Channel Evacuation Protocol for Spectrum-agile Networks," *IEEE DySPAN*, Apr. 2007.
[13] L.-C. Wang, C.-W. Wang, and C.-J. Chang, "Modeling and Analysis for Spectrum Handoffs in Cognitive Radio Networks," *IEEE Trans. Mobile Computing*, 2011.
[14] C.-W. Wang, L.-C. Wang, and F. Adachi, "Modeling and Analysis for Reactive-Decision Spectrum Handoff in Cognitive Radio Networks," *IEEE GLOBECOM '10*, Dec. 2010.
[15] —, "Optimal Admission Control in Cognitive Radio Networks with Sensing Errors," IEICE tech. rep., vol. 109, no. 440, Mar. 2010, pp. 491–96.

Biographies

LI-CHUN WANG [M'96, SM'06, F'11] (lichun@cc.nctu.edu.tw) received his B.S. degree from National Chiao Tung University, Taiwan, R.O.C., in 1986, his M.S. degree from National Taiwan University in 1988, and M.S. and Ph. D. degrees from Georgia Institute of Technology, Atlanta, in 1995 and 1996, respectively, all in electrical engineering. From 1996 to 2000 he was with AT&T Laboratories, where he was a senior technical staff member in the Wireless Communications Research Department. Since August 2000 he has been an associate professor in the Department of Communication Engineering of National Chiao Tung University, Taiwan. He was a co-recipient (with Gordon L. Stüber and Chin-Tau Lea) of the 1997 IEEE Jack Neubauer Best Paper Award. He has published over 150 journal and international conference papers. He was elected an IEEE Fellow in 2011 for his contributions in cellular architectures and radio resource management in wireless networks. He served as an Associate Editor for *IEEE Transactions on Wireless Communications* from 2001 to 2005, and as Guest Editor of Special Issues, on Mobile Computing and Networking for *IEEE Journal on Selected Areas in Communications* in 2005 and on Radio Resource Management and Protocol Engineering in Future IEEE Broadband Networks for *IEEE Wireless Communications* in 2006. He has eight U.S. patents.

CHUNG-WEI WANG [S'07] (hyper.cm91g@nctu.edu.tw) received his B.S. degree in electrical engineering from Tamkang University, Taipei, Taiwan, in 2003, and M.S. and Ph.D. degrees in applied mathematics and communication engineering from National Chiao Tung University in 2007 and 2010, respectively. From 2009 to 2010 he was also a visiting scholar at Tohoku University, Sendai, Japan. He was awarded student travel grants for IEEE ICC '09 and GLOBECOM '10. His current research interests include cross-layer optimization, MAC protocol design, and radio resource management in wireless sensor networks, ad hoc networks, and cognitive radio networks.

KAI-TEN FENG (ktfeng@mail.nctu.edu.tw) received his B.S. degree from National Taiwan University, Taipei, in 1992, his M.S. degree from the University of Michigan, Ann Arbor, in 1996, and his Ph.D. degree from the University of California, Berkeley, in 2000. Since August 2007 he has been with the Department of Electrical Engineering, National Chiao Tung University as an associate professor. He was an assistant professor with the same department between February 2003 and July 2007. He joined the Department of Electrical and Computer Engineering, University of California at Davis as a visiting professor between July 2009 and March 2010. He was with the OnStar Corp., a subsidiary of General Motors Corp., as an in-vehicle development manager senior technologist between 2000 and 2003, working on the design of future telematics platforms and in-vehicle networks. His current research interests include cooperative and cognitive networks, mobile ad hoc and sensor networks, embedded system design, wireless location technologies, and intelligent transportation systemsa. He received the Best Paper Award from the IEEE Vehicular Technology Conference Spring 2006, which ranked his paper first among the 615 accepted papers. He is also the recipient of the Outstanding Young Electrical Engineer Award in 2007 from the Chinese Institute of Electrical Engineering (CIEE). He has served on the Technical Program Committees of VTC, ICC, and WCNC.